BBC RD 1974/32

BBC • IBA

Engineering Research Report

The work covered by this report was undertaken by the BBC Research Department for the BBC and the IBA

UHF AMPLIFIERS: improvement of linearity by pre-correction

A.G. Lyner, B.Sc.
P. Shelswell, B.A.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION September 1974

BBC RD1974/32
UDC 621.375.4

UHF AMPLIFIERS: IMPROVEMENT OF LINEARITY BY PRE-CORRECTION

A.G. Lyner, B.Sc.
P. Shelswell, B.A.

Summary

The basic principle of pre-correction for improving amplifier linearity is discussed with some theory and the idea of pre-correction using an error signal injection system is introduced.

A description is given of a practical form of pre-corrector which was used successfully to double the output power of a transistor amplifier for a UHF television combined sound and vision signal.

Issued under the authority of

P. Rainger

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

September 1974

(RA-134)

UHF AMPLIFIERS: IMPROVEMENT OF LINEARITY BY PRE-CORRECTION

Section	Title	Page
	Summary	Title Page
1.	Symbols	1
2.	Introduction	1
3.	The theory of pre-correction	1
4.	Practical pre-correctors	2
5.	The construction of the chosen pre-corrector	3
	5.1. Signal divider and re-combiner unit	3
	5.2. Error function generator	4
	5.3. Complex attenuator	5
	5.4. Amplifier	5
6.	Bandwidth	6
7.	Experimental work	6
8.	Conclusions	7
	Appendix 1	7
	Appendix 2	8
	Appendix 3	8

UHF AMPLIFIERS: IMPROVEMENT OF LINEARITY BY PRE-CORRECTION

A.G. Lyner, B.Sc.
P. Shelswell, B.A.

1. Symbols

a_n	nth term of transfer characteristic for main amplifier
A	Gain of amplifiers in side chain of pre-corrector
b_m	mth term of pre-corrector transfer characteristic
C	Attenuation of complex attenuator
dBp	Decibels with respect to peak synchronising pulse power
$\epsilon(\nu)$	Error introduced by non-linearity in an amplifier. $\epsilon(\nu)$ is a function of ν
ϕ_n	Phase of nth component of a Fourier series
f_v	Vision carrier frequency
f_s	Sound carrier frequency
f_c	Colour subcarrier frequency
I_f	Level of intermodulation product at frequency f
λ	Wavelength of signal
l	Length of line in complex attenuator
m	Order of pre-correction term
n	Order of main amplifier transfer characteristic term
t	Time
ν_{in}	Instantaneous input voltage
ν_{out}	Instantaneous output voltage
ν_p	Pre-corrected signal
ω_n	Angular frequency
x, y	Two hypothetical signals

2. Introduction

Modern u.h.f. power transistors have been developed which are suitable for use in the output stages of low-power transmitters and transposers. Because the transistors have non-linear transfer characteristics, it is necessary to under-run them to reduce distortion unless a corrective system is employed. It is desirable therefore to have some means of compensation for the non-linear characteristic so that higher powers may be obtained from these amplifiers.

Negative feedback is not a practicable technique because power amplifiers have insufficient gain and excessive phase shift. The use of post correction (feed-forward) has been considered but for instrumental reasons it is not a suitable method at high power levels when efficiency is required. Envelope correction and correction at intermediate frequencies have been employed successfully in commercial equipment.

It is the purpose of this report to describe a system of pre-correction at u.h.f.

3. The theory of pre-correction

The distortion introduced into a signal by a transistor is predominantly amplitude distortion. Because of this the transfer characteristic may be represented by a power series.

$$\nu_{out} = \sum_{n=1}^{\infty} a_n \nu_{in}^n \qquad (1)$$

When operating at u.h.f., the coefficients a_n are frequency dependent but, initially, let it be assumed that they are effectively constant over the operating bandwidth.

An ideal pre-correction system would modify the input signal with a transfer law of the form

$$\nu_p = \sum_{n=1}^{\infty} b_m \nu_{in}^m \qquad (2)$$

giving an overall output

$$\nu_{out} = \sum_{n=1}^{\infty} a_n \nu_p^n \qquad (3)$$

$$= \sum_{n=1}^{\infty} \left[a_n \left(\sum_{m=1}^{\infty} b_m \nu_{in}^m \right)^n \right] \qquad (4)$$

Ideally this would reduce to a linear transfer characteristic:

$$\nu_{out} = a_1 b_1 \nu_{in} \qquad (5)$$

All amplifiers have an upper limit to their output power. Below this limiting level it is possible to choose a pre-correcting characteristic with coefficients b_m which are finite. Above this level, pre-correction is not possible. Additional limitations to the power output which can be accomplished using inverse law pre-correction include stability and the need for a very wide bandwidth for the pre-corrected signal.

An alternative way of considering pre-correction is to regard the pre-correcting process as the addition of an error signal to the input signal (Fig. 1). This error signal, for perfect pre-correction will be the same as the error signal produced with a feedback system if the loop gain were infinite and the feedback instantaneous. (Albeit allowing for discrepancies in gain between the two systems.)

This alternative approach has distinct practical advantages over the 'inverse-law' pre-corrector. The limitations to the production of a true inverse law are manifold and, even for a system involving purely static non-linearity, the actual measurement and adjustment of the coefficients of the law are difficult. In the case of a pre-corrector which injects an 'error-signal' however, the effects on the relevant parts of the signal spectrum may be studied and optimised

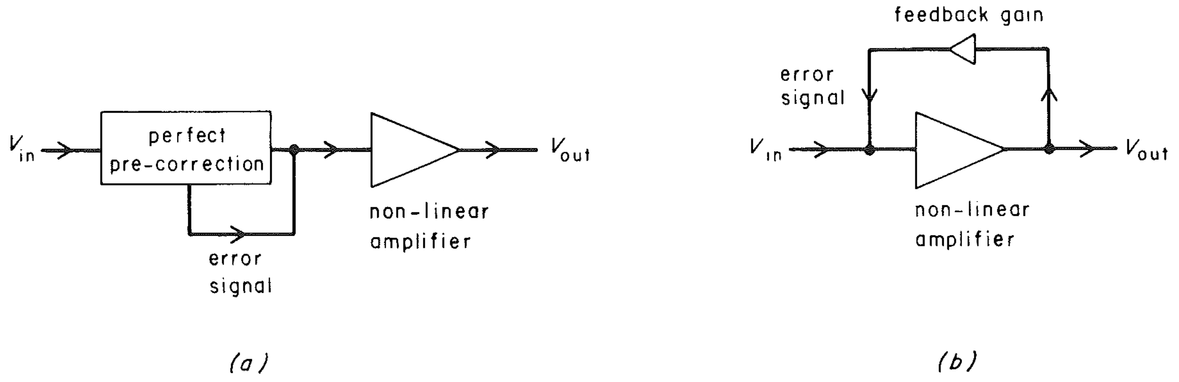

Fig. 1 - The equivalence of the error signal in pre-correction and feedback
(*a*) Pre-correction system (*b*) Feedback system

on their own and little attention need be paid to non-linear products which occur out of band as they can either be eliminated by filtering or ignored.

The amplifier distortion can be represented by an error signal $\epsilon(\nu)$ where

$$\epsilon(\nu) = \frac{1}{a_1} \sum_{n=1}^{\infty} a_n \nu_{in}^{n} \, , \qquad (6)$$

Appendix 1 shows that, when a pre-correcting circuit with transfer characteristic

$$\nu_{out} = \nu(_{in}) - \epsilon(\nu_{in}) \, , \qquad (7)$$

preceeds an amplifier with a transfer characteristic.

$$\nu_{out} = a_1 (\nu_{in} + \epsilon(\nu_{in})) \, , \qquad (8)$$

then it is possible to increase the power by 3·7 dB above that achievable with an uncorrected input whilst still keeping the relevant intermodulation products (i.p.) below −52 dB relative to peak sync power. The three-tone test signal* is used as an input in this case.

4. Practical pre-correctors

Two methods of producing a pre-corrector with the required characteristic are considered.

The first (Fig. 2(*a*)) consists of two parallel paths, one being a distortion free amplifier with 3 dB gain and the other being a model of the main amplifier with a transfer characteristic comparable with Equation (8). If these paths are fed from a power splitter and into a subtractive network, then the output will be the required function (Equation (8)).

The second method (Fig. 2(*b*)) takes a low level sample from the main signal path and splits it between two parallel paths, one path incorporating a suitable delay and the other a model of the main amplifier, attenuated to give 0 dB gain. When the signals from these paths are combined

---

* see Appendix 2

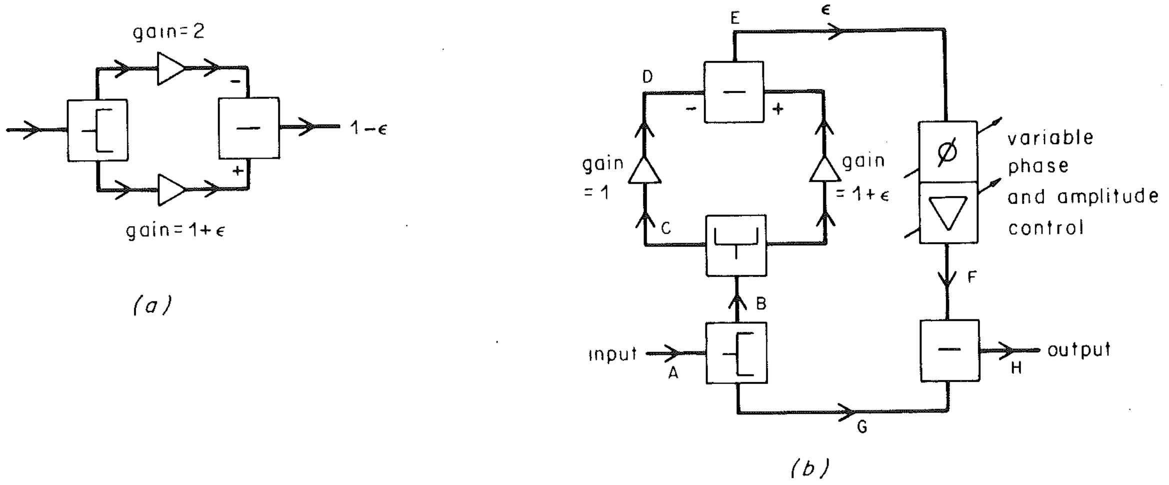

Fig. 2 - The two forms of pre-corrector considered initially

in a subtractor the result is the error function. This may be amplified and injected into the main path in the correct phase, to give the desired signal.

Consideration of the two projected methods of pre-correction and initial attempts at constructing both types yielded the following comparison of the basic merits of their respective designs. The differences between the two forms of pre-corrector lie in the ease of initial adjustment, their cost and the cost and ease of maintenance.

The first design operates at high power throughout and requires the use of expensive power transistors and associated bias circuits. This high cost is offset to some extent by the larger number of smaller components in the side chain of the second version, which is run at low power levels.

The initial adjustment of both units is critical, most of the controls being interactive. At the output of either pre-corrector the error signal must be of the correct amplitude and phase when compared with the main signal.

In the first system any control of the error signal will also affect the main signal. This interactive feature makes the system difficult to set up. The second system first isolates an error signal, then by a separate stage of control, suitably alters its amplitude and phase for injection into the main signal, thus simplifying the adjustment procedure.

The first design must be constructed in a single unit for reasons of stability. Assuming that service maintenance will be by replacement, the high cost of the unit is a disadvantage. It might also be considered wasteful to use high power devices in a unit which necessarily yields a low gain and will suffer from the poor reliability associated with such devices.

The second design, although more complex, should be more reliable, could be constructed in modular form and should be cheaper to build and maintain. It was decided that the second type of pre-corrector would be more suitable for use in service and it was chosen for more detailed study.

5. The construction of the chosen pre-corrector

The original pre-corrector was envisaged as an independent unit. It may, however, be incorporated with the amplifier with which it is associated. A block diagram of the chosen pre-corrector is shown in Fig. 3. It consists of four basic sections each of which can be constructed as a separate unit. The units are:

1. a signal divider and re-combining unit,
2. an error function generator,
3. a complex attenuator and
4. an amplifier.

Each unit will now be described in more detail.

5.1. Signal divider and re-combiner unit (Fig. 4)

In this unit the signal passes first through a 10 dB directional coupler, from which a sample is directed to the pre-corrector side chain. After this coupler there is a length of line through which the main signal passes to compensate for the delay through the side chain. This is followed by a second 10 dB coupler by means of which the error signal from the side chain is introduced into the main signal path.

The value of 10 dB for the couplers was chosen so that the loss in the main signal path was kept as low as possible without the requirement for excessive gain in the side chain (see Table 1).

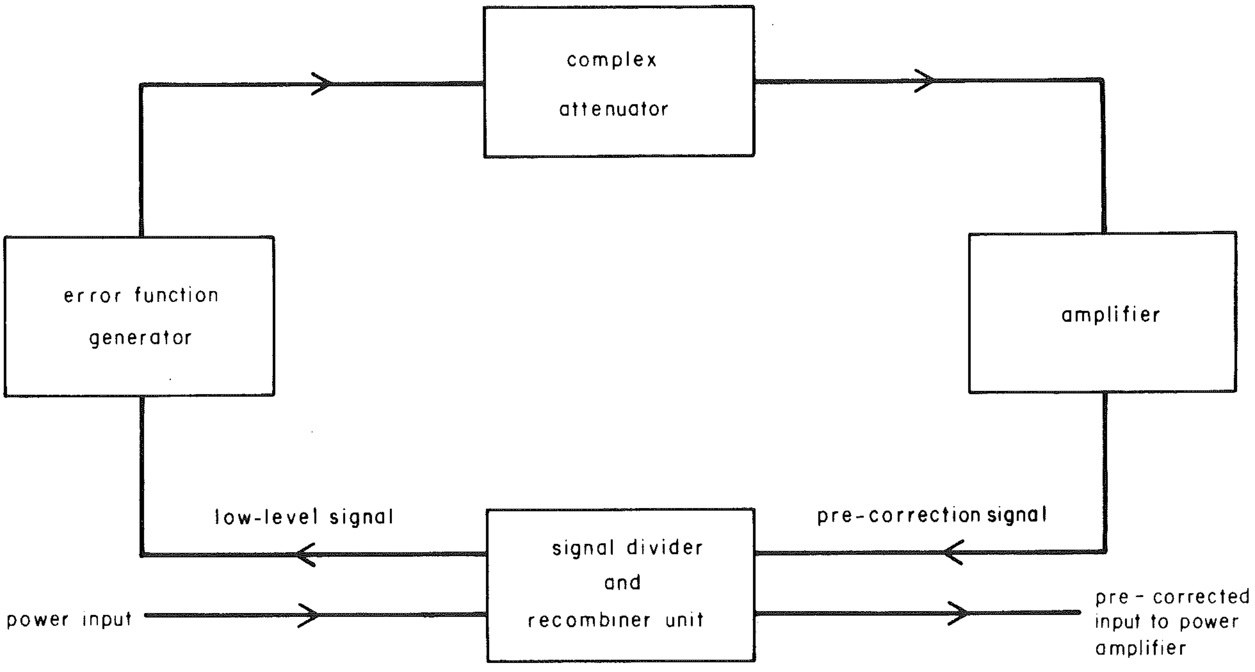

Fig. 3 - Schematic of pre-corrector units

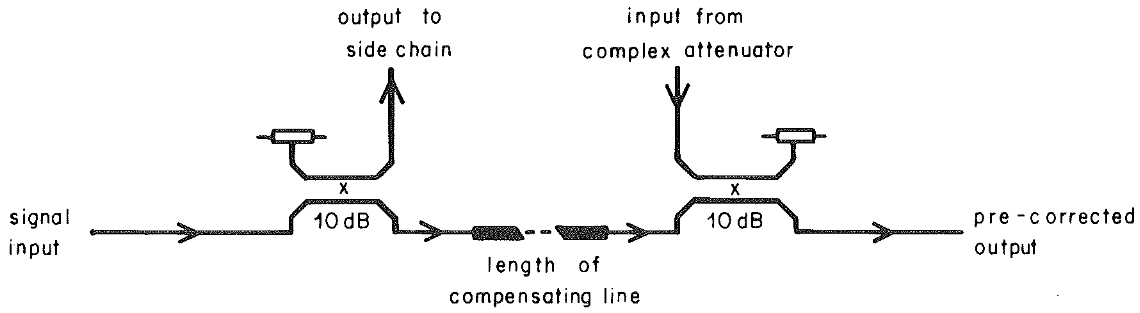

Fig. 4 - Signal divider and re-combiner unit

TABLE 1

Power levels with respect to input power at various points

	Place in Fig. 2(*b*)	Main signal dB w.r.t. input	Level of maximum intermodulation product dB w.r.t. input
Input	A	0	$-\infty$
Side chain output of first 10 dB coupler	B	-10	$-\infty$
Output of error function generator power splitter	C	-13	$-\infty$
After the amplifier in the error function generator	D	-13	-65
Output of error function generator	E	<-46	-68
After complex attenuator		$<(-46-C)$	$-68-C$
After amplifier	F	$<(-46-C+A)$	$-68-C+A$
Side chain input to output coupler	F	$<(-46-C+A)$	$-68-C+A$
Main chain input to output coupler	G	$-0{\cdot}5$	$-\infty$
Output	H	$-1{\cdot}0$	$-53=-78-C+A$

C is attenuation of complex attenuator
A is gain of amplifier

5.2. Error function generator

A circuit diagram of the error function generator is shown in Fig. 5. There are two parallel paths fed from a 3 dB directional coupler.

The first path contains a low-power model of the amplifier to be corrected followed by a variable resistive T attenuator which reduces the gain to unity. The second path is a length of line adjusted so that at the output of a combining directional coupler the two signal paths differ by 180°. The electrical length of the two paths are measured with a vector voltmeter by disconnecting and terminating each path in turn. Small adjustments in phase can be made by shunt capacitors in the output of the amplifier or on the parallel path, and compensating for amplitude changes with the T attenuator.

During the subtraction process in the error function generator it is necessary that the minuend and subtrahend are accurately in antiphase. If this is not the case the remainder may contain some of the original linear signal. The amount could be large enough to cause significant distortion when passed through the later stages of the pre-corrector. Appendix 3 shows the tolerances on phase and amplitude required to give various values of remainder. Ideally at low powers there would be no error introduced by the amplifier and therefore no remainder at all, but this is

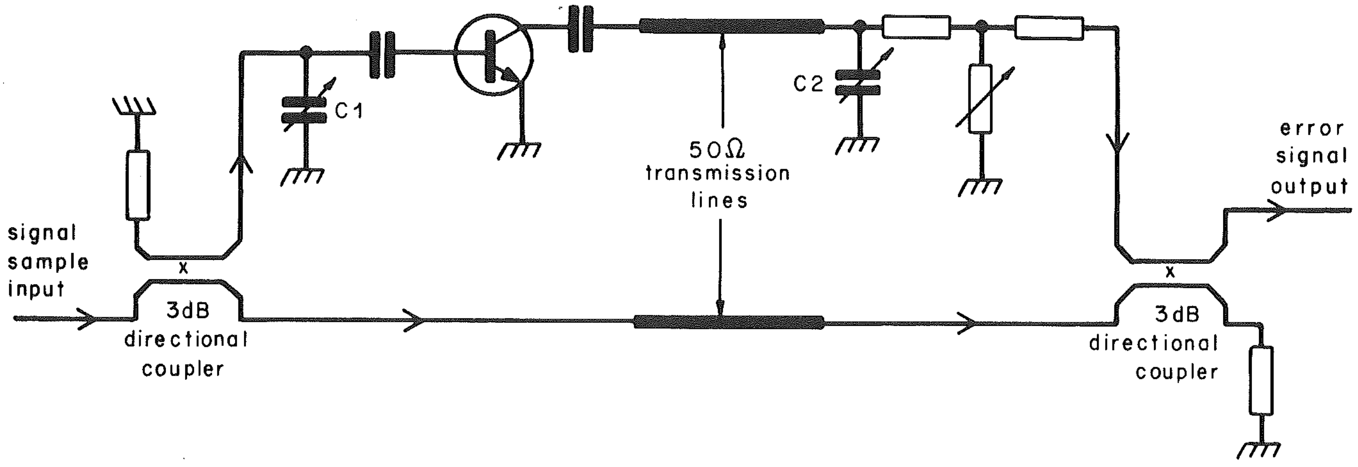

Fig. 5 - Signal paths in the error function generator

not achievable as unavoidable phase and amplitude differences across the band, imperfections in the directional couplers and stray fields, will lead to part of the original linear signal appearing as a residue. Allowing for these factors an acceptable residue is 30 dB below the level of the minuend. This implies a requirement for phase accuracies of the order of ±1° (line accuracies of about ±1 mm) and amplitude accuracies of about ±0·1 dB. Slight relaxations in construction tolerances may be allowed because of the fine adjustment allowed by C_2 and the variable attenuator.

The model amplifier must produce similar distortion to the main amplifier. For this reason it is necessary to use transistors rather than diodes as the distortion generators. UHF transistors are now available which belong to families of similar construction, but with various power handling capabilities. For instance, a commonly encountered series is the BLX 91, 92, 93 series. If the power amplifier uses BLX93 transistors, the model may use a single BLX91 or BLX92 provided the powers are adjusted so that the model is operated in a region of its characteristics comparable with that at which the main amplifier operates.

If the power amplifier uses transistors which do not belong to a given series the problem of modelling may be more difficult. At present there is no method of specifying all the parameters which are relevant so that the most suitable model may have to be selected by trial.

5.3. Complex attenuator

The complex attenuator consists of a 3 dB coupler (Fig. 6) with variable loads at the output ports, spaced l and (l + λ/8) from the coupler. The reflected waves form two quadrature vectors when combined at the output of the coupler. As the magnitude and sign of these vectors can be varied by altering the variable loads, the output signal has a variable phase and amplitude. The minimum insertion loss is between 3 dB and 6 dB because of loss in the loads and reflections to the input. In order to prevent these unwanted reflections from interfering with the operation of earlier stages, a matched buffer amplifier preceeds the attenuator. The output was also fed into a matched buffer amplifier to prevent later mismatches affecting the system.

5.4. Amplifier

Table 1 shows the power levels of the signal through the pre-corrector. It can be seen that for the intermodulation products to bear the same relationship to the main signal at the output of the model and at the output of the pre-corrector it is necessary to have A dB of amplification in the loop where

$$A = 25 + C$$

and C is the attenuation introduced by the complex attenuator.

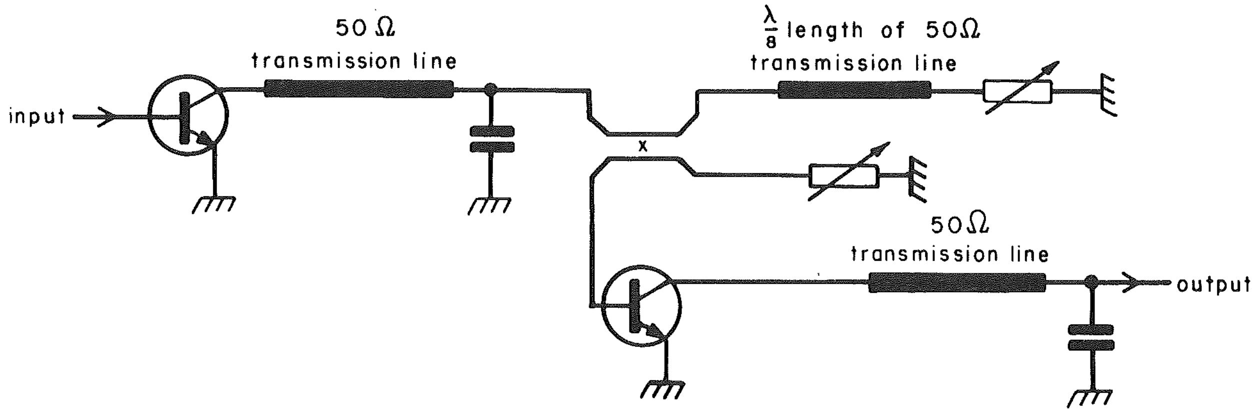

Fig. 6 - Complex attenuator and buffer stages

If the complex attenuator is used to adjust phase only, as is desirable, C will be about 6 dB.

i.e. $$A \simeq 31 \text{ dB}$$

For precise cancellation of distortions both amplitude and phase adjustments are necessary (see Appendix 3).

The maximum level which this amplifier must handle is −21 dB relative to input power. It must handle this level without distortion, otherwise the pre-correction signal will be affected. (N.B. The distortion components in this case account for a significant amount of the power through the amplifier.)

6. Bandwidth

The basic pre-corrector is envisaged as being a single-channel device, but capable of being adjusted during manufacture to operate in any one of a range of channels. When the units are assembled at the transmitter they will require a minimum of setting up.

It is possible to build amplifiers which can operate over a wide band by suitable adjustment of their matching and compensation. These can be built on printed circuit boards using microstrip techniques. Lengths of microstrip line can be adjusted by inserting lengths of coaxial cable.

This leaves the directional couplers as the only components whose frequency dependence may lead to different geometry requirements. As the signals from the main output ports are always in quadrature it is only the amplitude characteristic of the couplers which is important.

The directional couplers in the prototype were made of commercial screened twisted pair line, which is designed to give 0 dB output ratio when $\lambda/4$ long. There is a change in output ratio of less than −0·5 dB for a frequency spread of ±20% about the design frequency. This should enable one coupler to work over Band IV and one over Band V. Thus only two sizes of printed circuit board will be required for the error function generator and the complex attenuator, as the only directional couplers used in these have 0 dB output ratio. In the error function generator, any imbalance in the split signals may be compensated with the T attenuator. In the complex attenuator departure from 0 dB output ratio will result in increased insertion loss which can be allowed for in the following amplifier.

The directional couplers in the main power path have a coupling ratio of −10 dB. Commercial wireline is not available in a form which gives this coupling for lengths of $\lambda/4$.

A 10 dB coupler using available wireline has a variation in coupling ratio of the order of 1 dB over a 12% bandwidth. This high sensitivity to frequency variation makes it necessary to use two sizes of couplers in Band IV and three in Band V. This results in five different sizes of printed circuit board being necessary.

Other methods of making directional couplers exist which could be adapted to give wideband coupling values of −10 dB. The ease of making the board on which the −10 dB couplers are to be used is such that the cost of better couplers may not be warranted. This is a matter for further investigation.

Thus the method of pre-correction described may be used in both Bands IV and V, and requires few variations in printed circuit board geometry. Each unit must be adjusted to the correct channel during manufacture.

7. Experimental work

A system of the type described in Section 5 was built and tested. The amplifier to be pre-corrected used a BLX 93 transistor and the model in the error function generator was a BLX 91. Using optimum settings of the controls it was possible to increase output power for −52 dBp i.p.s from 0·5 dBW to +3·4 dBW. This represents an increase of 2·9 dB which compares favourably with the theoretical increase of 3·7 dB predicted in Appendix 1.

Investigations revealed that the main reasons for performance falling short of the theoretical maximum were break-through from the main signal path into the error function generator and failure to inject the error signal at the correct amplitude and phase into the main signal. Each of these factors hindered perfect vector subtraction of the third order products of the main amplifier.

The variation of i.p. level with output power for the practical pre-corrected amplifier was not of the form shown in Fig. A1 but varied about a straight line due, it was concluded, to low level third order products remaining at the output as a result of the pre-corrector imperfections described above.

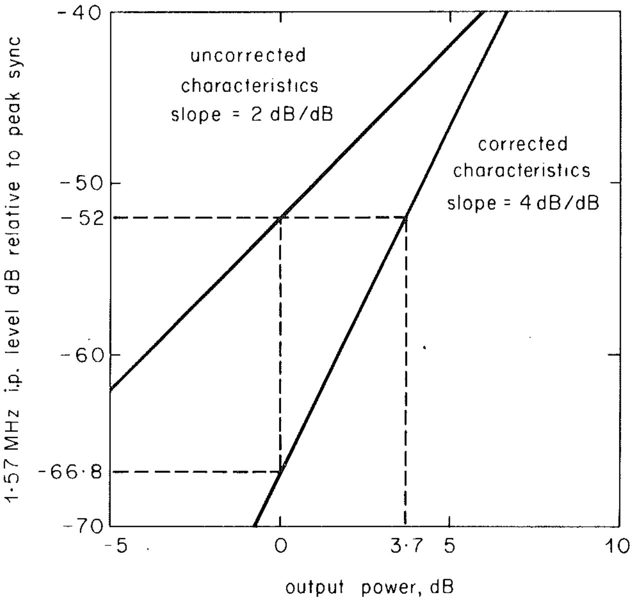

Fig. A1 - Relationship between distortion and power output for uncorrected and pre-corrected ideal non-linear amplifiers

Over the whole range of output powers below 3·5 dBW, however, i.p. levels did not exceed −52 dB w.r.t. peak sync.

8. Conclusions

A system of pre-correction has been described which enables greater output to be obtained from transistorised u.h.f. amplifiers with acceptable distortion. The method has been applied to the pre-correction of an amplifier, enabling more power to be obtained before intermodulation distortion exceeded −52 dB with respect to peak synchronising pulse power.

It is thought that this technique is suitable for use in conjunction with amplifiers in service.

Appendix 1

The mathematical theory of pre-correction

Assume an amplifier with a transfer characteristic

$$v_{out} = \sum_{n=1}^{\infty} a_n v_{in}^n \qquad \text{A1.1}$$

or

$$v_{out} = a_1 (v_{in} + \epsilon(v_{in})) \qquad \text{A1.2}$$

where

$$\epsilon(v_{in}) = \frac{1}{a_1} \sum_{n=2}^{\infty} a_n v_{in}^n \qquad \text{A1.3}$$

This can be preceeded by an amplifier with a characteristic

$$v_{out} = v_{in} - \epsilon(v_{in}) \qquad \text{A1.4}$$

giving an overall characteristic

$$v_{out} = a_1 v_{in} - a_1 \epsilon(v_{in}) + a_1 \epsilon(v_{in} + \epsilon(v_{in})) \qquad \text{A1.5}$$

which for a small error function ($\epsilon(v_{in})$) reduces to

$$v_{out} = a_1 \left(v_{in} + \frac{d\epsilon(v_{in})}{dv_{in}} \right) \qquad \text{A1.6}$$

If A1.6 is compared with A1.2 it can be seen that the error function of the pre-corrected system has been reduced by a factor $d\epsilon/dv_{in}$ which is always less than unity. Near saturation, other terms become appreciable in the series expansion of A1.5.

As the television signal is bandlimited, any even order terms in the error function will give rise to out-of-band distortions which are attenuated. Thus for practical purposes

$$a_n = 0 \quad \text{(n is even)} \qquad \text{A1.7}$$

Combining A1.1, A1.6 and A1.7 gives

$$v_{out} = a_1 (v_{in} + 3\left(\frac{a_3}{a_1}\right)^2 v_{in}^5 + \ldots) \qquad \text{A1.8}$$

as the output of the corrected system.

Applying the three-tone test signal (Appendix 2) to the original amplifier gives a series of intermodulation products. The most significant of these is a third order product at a frequency $f_v + f_s - f_{cs}$ and is at a level given by

$$I_{v+s-cs} = 20 \log_{10} \frac{3}{2}\left(\frac{a_3}{a_1}\right) g_v g_s g_{cs} \qquad \text{A1.9}$$

Thus for operation within the BBC specification where g_v, g_s and g_{cs} are the amplitudes of vision, sound and colour carriers respectively.

$$(I = -52\ \text{dB}) \qquad 20 \log_{10} \frac{a_3}{a_1} = -23{\cdot}5\ \text{dB} \qquad \text{A1.10}$$

The maximum intermodulation product from the corrected signal is a fifth order product also at a frequency $f_v + f_s - f_{cs}$ and for an input level giving uncorrected intermodulation products of −52 dB is at a level of

$$I_{v+s-cs} = -20 \log_{10} \left[\frac{45}{4}\left(\frac{a_3}{a_1}\right)^2 g_v g_s g_{cs} (g_v^2 + g_s^2) \right] \qquad \text{A1.11}$$

which will be −66·8 dB if A1.10 holds. If the input level is now increased, the intermodulation products, being fifth order, will increase by 5 dB for every 1 dB of input, a relative increase of 4 dB. Thus the input signal may be raised by 3·7 dB before the intermodulation products exceed −52 dB. See Fig. A1.

Appendix 2

The three-tone test

The standard distortion test used by the BBC is the three-tone test. Three c.w. tones represent vision carrier, sound carrier and colour subcarrier and are added at such power levels that their total when cophased is equal to the power of the synchronising pulse in the equivalent television signal.

The amplitudes and frequencies of the tones are listed in Table A2.1.

The present BBC specification states that all intermodulation products generated during the three-tone test and lying in the channel, must be at a level lower than −52 dBp.

TABLE A2.1

	Frequency (MHz w.r.t. vision carrier)	Level relative to peak sync (dB)
Vision carrier	0	−8 dB
Sound carrier	+6	−7 dB
Colour subcarrier	+4·433	−17 dB

Appendix 3

Errors in subtraction process

Assume two almost equal signals

$$x = \sum_{n=1}^{\infty} a_n \cos(\omega_n t + \phi_n)$$

$$y = \sum_{n=1}^{\infty} (a_n + \Delta a_n) \cos(\omega_n t + \phi_n + \Delta\phi_n)$$

Neglecting second orders of small quantities the difference signal is

$$x - y = \sum_{n=1}^{\infty} a_n \Delta a_n \cos(\omega_n t + \phi_n) - a_n \Delta\phi_n (\sin \omega_n t + \phi_n)$$

When this difference is split into amplitude and phase components the relative amplitude of each component depends on Δa_n and $\Delta\phi_n$ only.

$(x-y)/x$	$(a_n + \Delta a_n)/a_n$	or $\Delta\phi$
−10 dB	2·36 dB	18°*
−20 dB	0·82 dB	5·7°
−30 dB	0·24 dB	1·8°
−40 dB	0·08 dB	0·57°

* not accurate $\cos\Delta\phi \neq 1$
$\sin\Delta\phi \neq \Delta\phi$

SMW/AMM

Printed by BBC Research Department, Kingswood Warren, Tadworth, Surrey